(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
21 June 2001 (21.06.2001)

PCT

(10) International Publication Number
WO 01/44504 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/GB00/04746

(22) International Filing Date:
12 December 2000 (12.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9929720.2 17 December 1999 (17.12.1999) GB

(71) Applicant *(for all designated States except MG, US)*: **ASTRAZENECA AB** [SE/SE]; S-151 85 Södertalje (SE).

(71) Applicant *(for MG only)*: **ASTRAZENECA UK LIMITED** [GB/GB]; 15 Stanhope Gate, London WIY 6LN (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **FOX, Jayne, Catherine** [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). **HAQUE, Kemal** [GB/GB]; Alderley Park, Macclesfield, Cheshire SK10 4TG (GB). **LITTLE, Stephen** [GB/GB]; 1 Blacon Point Road, Chester, CH1 5LD (GB).

(74) Agent: **GAINEY, Laurence, David, Scott**; Astrazeneca, Global Intellectual Property, P.O. Box 272, Mereside, Alderley Park, Macclesfield, Cheshire SK10 4GR (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/44504 A2

(54) Title: DIAGNOSTIC METHOD

(57) Abstract: This invention concerns a method for detecting methylated CpG-containing nucleic acid by, contacting a nucleic acid containing specimen with an agent that modifies unmethylated cytosine; amplifying the nucleic acid region that contains the potential methylated CpG nucleic acid; separating and then reannealing the strands of the amplified DNA products; and detecting the presence of methylated nucleic acid based on duplex formation analysis. The method is particularly useful for prognosing, diagnosing or monitoring diseases associated with the methylation of nucleic acid, for example, tumours and certain developmental disorders.

DIAGNOSTIC METHOD

This invention describes a method for detecting methylated CpG-containing nucleic acid. This method may be applied to prognosing, diagnosing or monitoring diseases associated with the methylation of nucleic acid, for example, tumours and certain developmental disorders (Khan & Wood, *Curr. Opin. Neurol.*, 12: 149, 1999).

It is recognised that enzymatic modification of cytosine by the addition of a methyl group at the 5th position of the pyrimidine ring can occur in mammalian DNA *in vivo*. Such methylation occurs exclusively at the 5′ cytosine in CpG dinucleotide islands, so called because of their relative scarcity in the genomes of complex organisms such as man. The significance of DNA methylation in the control of gene expression is considered to be essential for normal growth and development (Antequera & Bird , in *"DNA Methylation: Molecular Biology and Biological Significance"*, Eds. Jost & Saluz, Birkhauser Verlag, Basel, p.16, 1993).

The silencing of one parental gene copy by methylation-mediated masking of the transcription start site, thus allowing selective expression of the remaining allele, is termed imprinting. The classical example of imprinting is the silencing of one of two X chromosomes in females. It is therefore not surprising that acquired methylation changes are now believed to be significant factors in the aetiology of a number of cancers (Laird & Jaenisch, *Annu. Rev. Genet.*, 30: 441, 1996), and inherited disorders such as Angelman/ Prader-Willi Syndrome (Khan & Wood, 1999, *supra*), Fragile X (Tapscott *et al.*, *Curr. Opin. Genet. Dev.*, 8:245, 1998), Duchenne muscular dystrophy (Yoshioka *et al.*, *Clin. Genet.*, 53:102, 1998), Autistic Disorder (Ashley-Koch *et al.*, *Am. J. Hum. Genet.*, 65(4 Suppl):A60) and carbamyl phosphate synthetase I deficiency (Strand *et al.*, *Am. J. Hum. Genet.*, 65(4 Suppl):A96).

Early methods for the detection of methylated cytosines (5meC) exploited the modified base's resistance to both sequence specific restriction endonuclease (Singer *et al.*, *Science* 203:1019, 1979) and chemical (i.e. hydrazine) cleavage (Church & Gilbert, *Proc. Natl. Acad. Sci. U S A*, 81:1991, 1984). After restriction or chemical cleavage, the genomic DNA is analysed by Southern blotting and hybridisation with a sequence-specific probe in order to detect uncleaved (5meC containing) DNA.

These approaches were refined by amplifying uncleaved DNA using polymerase chain reaction (PCR; US patent numbers 4,683,195 & 4,683,202 Roche) allowing the analysis of a

fraction of the amount of DNA required for Southern analysis, making it more applicable to the analysis of clinical samples (Pfeiffer *et al.*, *Science*, 246:810, 1989; Singer-Sam *et al.*, *Mol. Cell. Biol.*, 10:4987, 1990). The major shortfalls of the restriction endonuclease approach are: (a) the need to carefully control the efficiency of enzyme digestion or chemical cleavage as sub-optimal cleavage results in the persistence of intact sequences not containing 5meC, which may give rise to a detectable band after electrophoresis; and (b) the limitations imposed by the availability of CpG containing cleavage sites. In contrast, chemical cleavage procedures identify 5meC by the absence of a detectable band at a given position in a sequencing ladder. Background (non-specific) cleavage can confound accurate interpretation of the data.

More recent methods have exploited the resistance and sensitivity of 5meC and cytosine, respectively, to sodium bisulphite-mediated conversion to uracil (Frommer *et al.*, *Proc. Natl. Acad. Sci. USA*, 89:1827, 1992). This process occurs via the following chemical steps: sulphonation at the 6th position of the pyrimidine ring; hydrolytic deamination at position 4 to produce uracil sulphonate; and finally, desulphonation of the base. 5-meC is resistant to this conversion process, which is ultimately equivalent to the spontaneous deamination of cytosine which occurs *in vivo* and is a significant cause of mutation. Prior treatment of DNA with sodium bisulphite yields DNA in which unmethylated cytosines are converted to uracil, whereas methylated cytosines (5-methylcytosine) are resistant to this conversion and remain as cytosine. If sodium bisulphite treated DNA is then copied by a DNA polymerase using a normal (adenosine, thymidine, guanosine and cytidine) deoxynucleotide triphosphate pool, the products of the replication process will contain thymine at positions in the primary DNA sequence which previously contained cytosine, *i.e.* C to T transitions. Sequencing of the sodium bisulphite treated DNA will allow determination of the methylation status of the CpG sites present (Frommer *et al., supra*; Clark *et al., Nucleic Acids Res.*, 22:2990, 1994). The major disadvantage of this sequencing-based approach is the lack of sensitivity when detecting methylated sequences which are present as minority of the total population. Cloning fragments prior to sequencing greatly improves sensitivity but this procedure is labour intensive, technically difficult and expensive.

The recent development of a PCR based method has improved on previously reported approaches for methylation change detection (US patent number 5,786,146 Johns Hopkins University School of Medicine). This technique uses primers that are designed to

discriminate between 5meC and cytosine in DNA which has been subjected to sodium bisulphite conversion. The approach, therefore, relies on predicting the site(s) of, and designing primers specific for, the primary sequence change following bisulphite treatment. Thus, this method does not allow identification of methylation status in the sequence between the primer pair, unless amplicons are subjected to an additional sequencing step which may not detect under-represented changes. In addition, a methylation variant which was not *cis* and 3′ to the primer-specific change will not be detected because amplification itself is dependent on the sequence at the priming sites. In essence, specific primers must be designed for every CpG site to be investigated, even if the sites in question may only be tens of bases apart. Another disadvantage of this approach is the requirement for stringent optimisation and control of PCR reaction conditions and primer design in order to avoid the generation of false positive results due to the non-specific hybridisation and extension of primers.

The conversion of cytosine by sodium bisulphite has also permitted the exploitation of electrophoretic approaches for methylation analysis. The changes in the primary sequence induced by bisulphite treatment of differentially methylated DNA has allowed the discrimination of PCR products generated from methylated or unmethylated DNA by single strand conformation polymorphism (SSCP) analysis (Bianco *et al.*, *Hum. Mutat.*, 14:289, 1999; Maekawa *et al.*, *Biochem. Biophys. Res. Commun.*, 262:671, 1999; Burri & Chaubert, *BioTechniques*, 26:232, 1999). Amplicons which differ by as little as a single base (generated by the presence of 5meC or cytosine) may form differing conformers and migrate differently during electrophoresis. Alternatively, PCR products which have been generated from differentially methylated DNA will vary in their thermal stability due to their different GC contents, and as a consequence can be separated by denaturing gradient gel electrophoresis (DGGE; Aggerholm *et al.*, *Cancer Res.*, 59:436, 1999). An advantage of both electrophoretic approaches is their potential to interrogate more than one CpG site at one time. A major drawback of these strategies is the operational requirement for highly reproducible optimised gel conditions in order to identify all of the sequence variants that could be generated by bisulphite conversion. Slight variations in reproducibility may significantly reduce the resolving power of electrophoresis resulting in a significant decrease in sensitivity. This potential problem will become particularly acute if the methylated sequences constitute only a fraction of the total population in the samples under investigation.

Aberrant methylation changes are of increasing interest to the cancer researcher and are now considered to be a mechanism (in addition to gene mutation, amplification and loss) of deregulation of gene expression in cancer cells (Laird & Jaenisch, 1996, *supra*). Loss or increase in gene expression are important criteria when attempting to ascribe tumour suppressor or oncogene function to a gene (Whang *et al.*, *Proc. Natl. Acad. Sci. U S A*, 95:5246, 1998). This being the case, together with the limitations of current methods, it becomes apparent that the development of improved approaches capable of detecting acquired methylation changes (when present in a minority of cells in a biological sample), is highly desirable. Table 1 provides a non-comprehensive list of genes whose methylation status may be an aetiological factor in carcinogenesis.

Tabel 1 Hypermethylation of Putative Tumour Suppressor Genes in Human Cancers

Gene	**Chromosomal Locus**	**Tumour Type**
RB1	13q14	retinoblastoma
VHL	3p26-25	renal cancers
CDKN2A	9p21	melanoma & others
CDKN2B	9p21	haematologic malignancies
MLH1	3p21	colorectal cancer
APC	5q21	colorectal cancer
PTEN/MMAC1	10q23	prostate cancers & others
BRCA1	17q21	breast and ovarian cancers
M6P/IGF2r	6q26	breast cancer
WT1	11p13	paediatric kidney cancer

According to a first aspect of the invention there is provided a method for detecting a methylated CpG-containing nucleic acid comprising,

a) contacting a potential 5meC containing nucleic acid test sample with an agent that selectively deaminates unmethylated cytosine to uracil;

b) amplifying the region(s) from the nucleic acid sample that contains the potential site(s) of methylation;

c) separating the strands of the amplicons;

d) allowing the separated strands to reanneal and form duplex molecules; and.

e) detecting mismatch-containing heteroduplexes arising from the annealing of strands generated from methylated and unmethylated nucleic acid.

According to a further aspect there is provided a method for detecting a methylated CpG-containing nucleic acid comprising:

a) contacting a nucleic acid containing specimen with an agent that modifies unmethylated cytosine;

b) amplifying the nucleic acid region that contains the potential methylated CpG nucleic acid;

c) separating the strands of the amplified DNA products;

d) allowing the separated strands to reanneal to form duplex molecules; and

e) detecting the presence of a methylated nucleic acid based on aberrant duplex formation or the type(s) of duplexes formed.

The method of the invention calls for modification of the test nucleic acid by sodium bisulphite or comparable agent(s) which converts unmethylated cytosines (but not methylated cytosines) to uracil, subsequent amplification of the treated nucleic acid using an amplification reaction such as the polymerase chain reaction (PCR), followed by separation and reanneal of the amplified nucleic acid strands and duplex analysis of the treated nucleic acid to determine methylation status. The principal agent used in the art for selectively deaminating unmethylated cytosine to uracil is sodium bisulphite.

The specimen containing the sample nucleic acid is preferably one isolated from an animal, preferably a human tissue or fluid sample. Such a sample may conveniently be from a solid tissue, such as from a tumour or tumour margin, or other biopsy sample, or from a stool sample or bodily fluid sample (such as, sputum, saliva, blood, semen, urine and the like). The sample may be fresh or one preserved by for example, freezing, formalin, or other tissue fixation methods, and may then optionally be embedded in paraffin or the like. The nucleic acid may be present within the sample or may be in a solution extracted therefrom. If necessary, the nucleic acid can be extracted from the biological specimen using a variety of techniques, for example, as described by Maniatis *et al.* (*"Molecular Cloning: A Laboratory Manual"*, Cold Spring Harbour, NY, pp280-281, 1982). Any of the commercially available nucleic acid extraction/purification kits can also be used. The nucleic acid may be also be pre-amplified using for example the PCR prior to analysis. In a preferred embodiment the nucleic

acid in the test sample is isolated from the biological sample. In another preferred embodiment the test nucleic acid is quantified before subsequent analysis.

By way of illustrating a suitable method, the test nucleic acid is preferably quantified by any convenient means. The DNA is then denatured by, for example treatment with NaOH, prior to incubation with the modifying agent sodium bisulphite. Modified DNA is then purified in order to desalt using any convenient method/commercially available kit. Modification is completed by treatment with NaOH before ethanol precipitation, and resuspension in water or Tris buffer (Frommer *et al., supra*; Clark *et al.,* 1994; Herman *et al.*, *Proc. Natl. Acad. Sci. USA*, 93:9821,1996). Following cytosine conversion, the target region under study is amplified using a suitable target nucleic acid amplification procedure such as polymerase chain reaction (PCR). Most preferably, the PCR primers bind outside the site or zone or region of methylation. Other suitable amplification methods are identified below. Following this the strands of the amplicons are separated by, for example, heating to 95°C for 5 minutes and then allowed to reanneal by bringing the temperature down to room temperature over 20 minutes. Aberrant duplex formation indicative of the presence of methylated CpG, is then determined by DHPLC analysis.

The polymerase chain reaction (PCR) procedure can be used to amplify specific nucleic acid sequences through a series of iterative steps including denaturation, annealing of sequence specific oligonucleotide primers, and extension of the primers with DNA polymerase (US patent numbers 4,683,195 & 4,683,202). Other known nucleic acid amplification procedures include transcription-based amplification (Kwoh *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:1173, 1989), Strand Displacement Amplification (SDA; Walker *et al.*, *Nucleic Acids Res.*, 20:1691, 1992), rolling circle amplification (Lizardi *et al., Nat. Genet.*, 19:225, 1998) and Nucleic Acid Sequence Based Amplification (NASBA; Compton, *Nature*, 350:91, 1991)

The PCR is the most preferred and widely used nucleic acid amplification procedure. When using PCR, it may be desirable to design two sets (pairs) of PCR primers; the first set complementary to the predicted DNA sequence after sodium bisulphite treatment, *i.e.* adenine should be used at all positions in the primer oligonucleotide which will lie anti-parallel to, and therefore base pair with, a position in the template DNA containing uracil; and, the second set of primer oligonucleotides (WT sequence) should contain guanine at positions where the first set contained adenine and is thus complementary to cytosine in the untreated DNA. The use of

both primer sets, coupled with quantitative methods of DNA amplification would allow control for the specificity of the primers and efficiency of the bisulphite C to U conversion process. Amplifying untreated DNA with WT primers would allow the identification of mismatch heteroduplexes which have been generated as a result of primary sequence heterogeneity (*e.g.* single nucleotide polymorphisms or somatic mutations) as opposed to differences in methylation of CpG sites under investigation. Thus, in a preferred embodiment polymerase chain reaction amplification is performed on treated and/or untreated sample DNA using primers capable of annealing to untreated WT and/or modified nucleic acid. Whilst it is possible to interrogate both sense and anti-sense strands simultaneously in the same reaction, it is preferred that only one of the strands be interrogated in a single reaction. This is because, as the complementarity of the treated strands becomes diminished, potential non-5meC derived duplexes that might form would increase the complexity of subsequent duplex analysis. Unless it is known *a priori* whether the sense or anti-sense strand is methylated, primers specific for each strand (post-bisulphite treatment) should be designed.

When amplifying the test region it is preferable to select primer binding sites which do not overlie potential CpG methylation sites. If sequence constraints dictate that primers overlie such sites, mismatches will result if the priming sequence in the sample nucleic acid is differentially methylated, which may in turn have an inhibitory effect on amplification. Potential inhibition of PCR could be avoided if primers are designed to contain base analogues (eg nitroindole, nitropyrrole and inosine) at sites which will base-pair with a cytosine which may be differentially methylated. The analogues will base-pair with both cytosine and uracil and thus priming will be unaffected by the methylation status of the CpG site(s) in question.

Amplification reaction conditions can be optimised to account for melting temperature of the primer/template hybrid, the size of amplification product to be generated, integrity and quantity of template nucleic acid *etc.*. The preferred length of target DNA region is between 30 and 1000 bp, more preferably between 50 and 500, most preferably between 100 and 400 bp. Because clinical tissue specimens such as paraffin embedded tissue biopsies are often found to be partially degraded, it can be technically difficult to amplify large fragments. Consequently, amplified PCR products from such samples should preferably be about 100-250bp in size. However, future development of improved clinical procedures for sampling and

preserving tissue specimens may be expected to relax this size restriction and there is believed to be no theoretical limit to the preferred size range of the amplified products.

Potentially numerous target regions could be interrogated simultaneously in a multiplex PCR. The reaction conditions adopted will have to be a compromise based on the individual optimum conditions for each reaction. The precise conditions to be adopted can be determined by the person skilled in the art. In one embodiment of the method, a single primer pair could be used in the amplification of bisulphite treated and untreated DNA.

The amplified products are subjected to heteroduplex analysis which will permit the detection of sequences generated from 5meC containing DNA. Sites in the original template sequences which were either methylated or unmethylated (such heterogeneity being characteristic of clinical samples) will give rise to cytosine or thymine, respectively. If the sense and anti-sense strands of these amplicons are separated and reannealed, in addition to perfectly matched homoduplexes, heteroduplex molecules which contain either C:T or G:A single base mismatches will be formed. The presence of mismatched heteroduplex is indicative of methylation differences in the target sequence in the original nucleic acid sample.

Duplexes are formed by separating and reannealing the amplified regions of interest. Typically, this can be achieved by heat denaturing a solution containing the amplified regions of interest, followed by cooling to allow the melted DNA strands to reanneal. However, it is not intended that the method of the invention should be restricted to thermal techniques, and any convenient method for duplex separation (i.e alkali denaturation) and formation may be used.

Heat denaturation is conveniently carried out by subjecting the nucleic acid sample to temperatures around 95°C, for example between 92°C and 100°C, for a duration sufficient to ensure strand separation, nominally at least 1 minute, usually between 2 and 10 minutes. Annealing is generally carried out by allowing the temperature of the denatured solution to drop to 37°C over a period of 1 to 2 min. A more gradual cooling rate of 1-4°C per minute may be preferred. The optimum denaturing temperature and annealing rate will depend on the duplex composition and length. The optimum temperatures and times required to ensure denaturation and annealing can be determined by the person skilled in the art.

The method assumes that the relative number of heteroduplexes and homoduplexes formed will depend only on the relative frequency of the two alleles. In practice, it is possible

that homoduplexes may form in preference to heteroduplexes because of the greater binding affinity of the perfectly matched sequences. The introduction of a thermostable region into the product so that any potential annealing bias due to the mismatch(es) will be overpowered by the thermostable region should ensure random annealing of nucleic acid molecules. Therefore, in a preferred aspect of the invention the method will incorporate a technique to promote the formation of stable heteroduplexes. In a preferred embodiment of the invention the amplification primers will have a thermostable region incorporated therein to promote the formation of stable heteroduplexes. In a more preferred embodiment, the thermostable region is a GC clamping sequence that has been incorporated into the design of the PCR primers used to amplify the nucleic acid regions of interest (see for example, Hayes *et al.*, *Diagn. Mol. Pathol.*, 8:2, 1999)

Heteroduplexes may be detected by any convenient method, for example, physical, enzymatic or chemical mismatch cleavage, or mismatch binding. In a preferred aspect of the invention, heteroduplexes are detected by binding to prokaryotic or eukaryotic mismatch binding proteins. An example is *MutS*, a mismatch binding protein isolated from *E.coli*, which recognises regions of double-stranded DNA containing a single mismatched base pair (Wagner *et al.*, *Nucleic Acids Res.*, 23, 3944, 1995). *MutS* is allowed to bind to the heteroduplexes and bound heteroduplex/*MutS* complexes are removed from the reaction mixture using, for example, powdered nitrocellulose. A convenient alternative is to use *MutS* conjugated to magnetic beads, allowing bound heteroduplexes to be removed from the reaction mixture with a magnet. *MutS* may also be conjugated to biotin and the bound heteroduplexes removed from the mixture using streptavidin coated beads. Bound heteroduplexes may be detected using, for example, intercalation with a fluorescent dye by pre-labelling the primers used for amplification with a fluorophore, or directly incorporating the fluorophore into the product of amplification. Alternatively, amplicons can be labelled with haptens such as digoxigenin and dinitrophenol.

In another preferred aspect of the invention mammalian or bacterial endonucleases are used to recognise and cleave the heteroduplexes at mismatched bases (see US Patent No.5,824,4710). Examples of preferred enzymes include bacteriophage resolvases such as T4 endonuclease VII or T7 endonuclease I. The detection of cleavage products would be indicative of the presence of mismatch containing heteroduplexes.

The most preferred method of detecting heteroduplex molecules involves physical separation, such as achieved by chromatography or electrophoresis. Suitable examples include, denaturing high performance liquid chromatography (DHPLC) and chemical or temperature denaturing electrophoresis. Denaturing HPLC, a chromatographic technique capable of separating heteroduplex and homoduplex DNA molecules in a mixture, is as described in United States Patent No. 5,795,976 and is incorporated herein by reference. The reader is referred thereto for information on the working of DHPLC. The duplex mixture is applied to a stationary reverse-phase support and the homo and heteroduplex molecules are eluted (under thermal or chemical conditions capable of partially denaturing heteroduplexes) with a mobile phase containing an ion-pairing reagent (*e.g.* triethylammonium acetate; TEAA) and an organic solvent (*e.g.* acetonitrile; AcN). DHPLC can also allow the direct quantitation of relative homoduplex and heteroduplex concentrations by the detection of ultraviolet absorbance or fluorescent emission of/from the separated species. The area under the absorbance/emission peak is proportional to the amount of product which therefore allows quantitative assessment of heteroduplexes. DHPLC is described in Liu W *et al.* (Nucleic Acids Research. 26:1396-1400, 1998 and O'Donovan MC *et al.* Genomics. 52:44-49, 1998).

In the situation where abnormal methylation is under investigation in a given gene sequence the selection of appropriate control DNA sources may be required. For example, when investigating the role of methylation in a tumour sample it may be necessary to select a normal tissue source as reference. However, an alternative approach may be to use heteroduplex migration/elution (depending on the physical separation method employed for duplex analysis) pattern differences of bisulphite modified to unmodified DNA. For example, when the normal methylation status of a particular gene in a specific tissue type is known, then the difference between bisulphite treated and untreated duplex migration may be sufficient to identify methylation pattern differences without further reference to control DNA profiles. The choice of appropriate controls and methodological approaches will depend on the specific application intended.

Sensitive detection techniques such as laser induced fluorescence (LIF) of labelled duplexes, will allow the detection of methylation pattern differences in a small fraction of the sample under analysis. For example, LIF detection of DHPLC separated duplexes is stated to be sensitive enough as to allow the detection of duplex species which are present at a level of 0.2% of the total population (Hecker *et al., Anal. Biochem.*, 272:156, 1999; PCT WO

99/19514). This level of sensitivity would facilitate the detection of aberrantly methylated cancer cell-derived nucleic acid sequences when they constitute only a fraction of the total population in a clinical sample. In a preferred embodiment the duplex analysis is performed by DHPLC coupled with LIF.

Thus according to a further aspect of the invention there is provided a method for detecting a methylated CpG-containing nucleic acid comprising:

a) contacting a nucleic acid containing specimen with an agent that modifies unmethylated cytosine;

b) amplifying from within the nucleic acid sample the region that contains the potential methylated CpG nucleic acid;

c) separating the strands of the amplified DNA products;

d) allowing the separated strands to reanneal to form duplex molecules; and

e) analysing the duplex products by dHPLC, optionally coupled to LIF.

Some applications of the instant method for methylation pattern difference detection are identified below:

1. The detection of methylation changes in genes in order to diagnose an imprinting-related disorder such as Angelman or Prader-Willie syndromes.

2. The identification and study of methylation as a cause of aberrant transcription and its role in the development of neoplasia. This includes the use of this method to help to ascribe oncogene or tumour suppressor function to novel or known gene(s).

3. The sensitive detection of a sub-population of cells containing methylation pattern(s) uncharacteristic of DNA normally derived from that tissue type. Specific uses of this may include the identification of pre-malignant cells in a tissue sample.

4. The methylation pattern of a primary tumour may be used as a fingerprint to provide a method to characterise the tissue origin of metastases. This information can be an important factor in defining an appropriate treatment regime and/or determining disease prognosis.

5. The analysis of methylation patterns in the study of developmental control and regulation.

6. The identification of novel, imprinted genes.

The above disclosure generally describes the above invention and will be illustrated, but not limited, by reference to the following Example and Figures in which:

Figure 1 shows the conversion of 5meC to uracil and the subsequent amplification of the target sequence. The amplicons thus generated are subjected to heteroduplex analysis. Untreated DNA is also amplified as a control.

Figure 2 demonstrates the formation of mismatch heteroduplexes when sequence variant amplicons generated from sodium bisulphite treated template DNA are denatured and reannealed. Control (untreated) DNA amplicons will only form homoduplexes.

Figure 3 shows DHPLC separation of homoduplex and mismatch heteroduplex species.

Example 1: Methylation Analysis Using Physical Heteroduplex Detection Method

Sodium Bisulphite Treatment

Purified genomic DNA (1μg) is denatured in 0.2M NaOH for 10-15 minutes at 37°C (final volume 20μl). For lower amounts of DNA (*e.g.* 5-500ng), salmon sperm DNA is added to give a final quantity of 1μg prior to denaturation. 278μl of 4.8M sodium bisulphite (pH 5) and 2μl of 100mM hydroquinone (both freshly prepared) are added and the solution mixed by inversion prior to overlaying with mineral oil and incubation at 55°C for 5 hours. The DNA is then purified using a suitable kit, *e.g.* Wizard DNA Purification (Promega) or Qiaquick Purification (Qiagen) systems according to manufaturer's instructions, and recovered in 50μl H_2O. 5.5μl of 3M NaOH hydroxide is added, and the modification process completed by incubation at 37°C for 15 minutes. The DNA is recovered by ethanol precipitation and resuspended in 10-50μl H_2O.

PCR Amplification of the DNA Region Containing the CpG Sites

PCR primers pairs which (a) are specific for the sense-strand of the target sequence, (b) flank the CpG sites of interest, and (c) bind to either bisulphite treated or untreated DNA (control primers) are designed and synthesised. DNA from [1] (5-500ng) is amplified (see Figure 1) in a PCR reaction designed to generate a 100-500bp amplicon (typically 1-2 units *Taq* or other thermostable polymerse; 30-40 amplification cycles comprising 94-95°C for 30-60 seconds (denature), 50-65°C for 30-60 seconds (anneal), 72°C for 30-60 seconds (primer extension)).

Formation of Heteroduplexes

PCR products from [2] are denatured by heating at 94°C for 1-5 minutes. Sequence variants, which have been generated by amplifying sodium bisulphite treated, differentially methylated DNA, are allowed to form mismatch containing heteroduplexes by cooling to 37°C over a period of 1-2 minutes (see Figure 2). Although it has been reported that such conditions

favour the formation of mis-match containing heteroduplexes over matched homoduplexes (Oka *et al.*, *Nucleic Acids Res.*, 22:1541), it may be preferable to allow annealing to occur over a longer period of time (up to 1 hour or more) due to the actual annealing characteristics of the sequence under investigation.

Detection of Mismatch Heteroduplexes

A Wave™ Fragment Analysis System (Transgenomic Inc.) can be used for analysis. The sample from [3] is applied to a DHPLC column. Typically, chromatographic conditions for separation of homo and heteroduplexes employ a column temperature of 50-60°C and a 0.1M triethylamine acetate (pH7) mobile phase containing 10% acetonitrile (increasing to 25% over a period of 5-8 minutes; flow rate ~0.75-0.9 ml/minute). Chromatographic peaks (detected by absorbance at 260nm) corresponding to homoduplexes and heteroduplexes are resolved (see Figure 3). The presence of a heteroduplex peak (or peaks - depending on the number of CpG sites flanked by the amplification primers) in bisulphite-treated, as opposed to the untreated reaction, is indicative of 5meC in the original sample. The absence of a heteroduplex peak in the treated and untreated reactions indicate that the original sample was unmethylated. Furthermore, the ratio of heteroduplex to homoduplex peak area will be directly related to the proportion of methylated and unmethylated target sequences present in the original DNA sample. Therefore this approach has the potential to facilitate both qualitative and quantitative analysis of gene methylation.

Increasing the Sensitivity of Detection

The combination of laser induced fluorescence (LIF) and DHPLC technology has considerably increased the sensitivity of duplex detection. DHPLC-LIF of labelled duplexes has been demonstrated to allow the detection of duplex species which are present at a level of 0.2% of the total population (Hecker *et al.*, 1999; PCT WO 99/19514). This level of sensitivity would facilitate the detection of aberrantly methylated cancer cell-derived nucleic acid sequences when they constitute only a fraction of the total population in a clinical sample. The method is performed as described above in steps [1] to [4], except amplification primers are 5′-labelled with a fluorogenic moiety such as 6-carboxyfluoroscein (FAM) or tetrachloro6-carboxyfluoroscein (TET).

Claims:

1. A method for detecting a methylated CpG-containing nucleic acid comprising:

a) contacting a nucleic acid containing specimen with an agent that modifies unmethylated cytosine;

b) amplifying the nucleic acid region that contains the potential methylated CpG nucleic acid;

c) separating the strands of the amplified DNA products;

d) allowing the separated strands to reanneal to form duplex molecules; and

e) detecting the presence of a methylated nucleic acid based on the type(s) of duplexes formed.

2. The method as claimed in claim 1, wherein step b) amplification involves polymerase chain reaction (PCR) using primers that bind either side of the site(s) of methylation.

3. The method as claimed in claim 1, wherein step b) amplification involves polymerase chain reaction using primers which contain a base analogue capable of base pairing with uracil and cytosine with comparable efficiency at the complementary site position of a potential methylated site.

4. The method as claimed in claim 2 or 3, wherein one or more of the PCR primers have a thermostable region incorporated therein to promote the formation of stable heteroduplexes.

5. The method as claimed in claim 4, wherein the thermostable region is a GC clamping sequence that has been incorporated into the design of the PCR primers used to amplify the nucleic acid region of interest.

6. The method as claimed in any of claims 2 - 5, wherein two sets of PCR primers are used, one set complementary to the predicted DNA sequence after sodium bisulphite treatment and the other set complementary to predicted untreated DNA sequence.

7. The method as claimed in any of the preceding claims, wherein amplification is also carried out on untreated DNA as control.

8. The method as claimed in any of the preceding claims wherein the agent that modifies unmethylated cytosine is sodium bisulphite.

9. The method as claimed in any of the preceding claims, wherein the test nucleic acid is quantitated prior to step (a).

10. The method as claimed in any of the preceding claims, wherein duplex formation in step (e) is determined by denaturing high performance liquid chromatography (DHPLC), optionally coupled to laser induced fluorescence (LIF).

11. The method as claimed in any of claims 1 - 9, wherein duplex formation in step (e) is determined by enzymatic or chemical cleavage of the mismatch duplex products formed in step(d) and detecting said products by electrophoresis or ion-pair reverse phase HPLC, optionally coupled to laser induced fluorescence (LIF).

12. The method as claimed in any of claims 1 - 9, wherein duplex formation in step (e) is determined by gel electrophoresis.

13. A method for detecting a methylated CpG-containing nucleic acid comprising:

a) contacting a nucleic acid containing specimen with an agent that modifies unmethylated cytosine;

b) amplifying from within the nucleic acid sample the region that contains the potential methylated CpG nucleic acid;

c) separating the strands of the amplified DNA products;

d) allowing the separated strands to reanneal to form duplex molecules; and

e) analysing the duplex products by denaturing high performance liquid chromatography (DHPLC), optionally coupled to laser induced fluorescence (LIF).

14. A method for detecting a methylated CpG-containing nucleic acid comprising:

a) contacting a nucleic acid containing specimen with an agent that modifies unmethylated cytosine;

b) polymerase chain reaction amplification of the methylated CpG-containing nucleic acid using primers that bind outside the methylated CpG-containing nucleic acid region of interest;

c) separating the strands of the amplified DNA products;

d) allowing the separated strands to reanneal to form duplex molecules; and

e) analysing the duplex products by denaturing high performance liquid chromatography (DHPLC), optionally coupled to laser induced fluorescence (LIF).

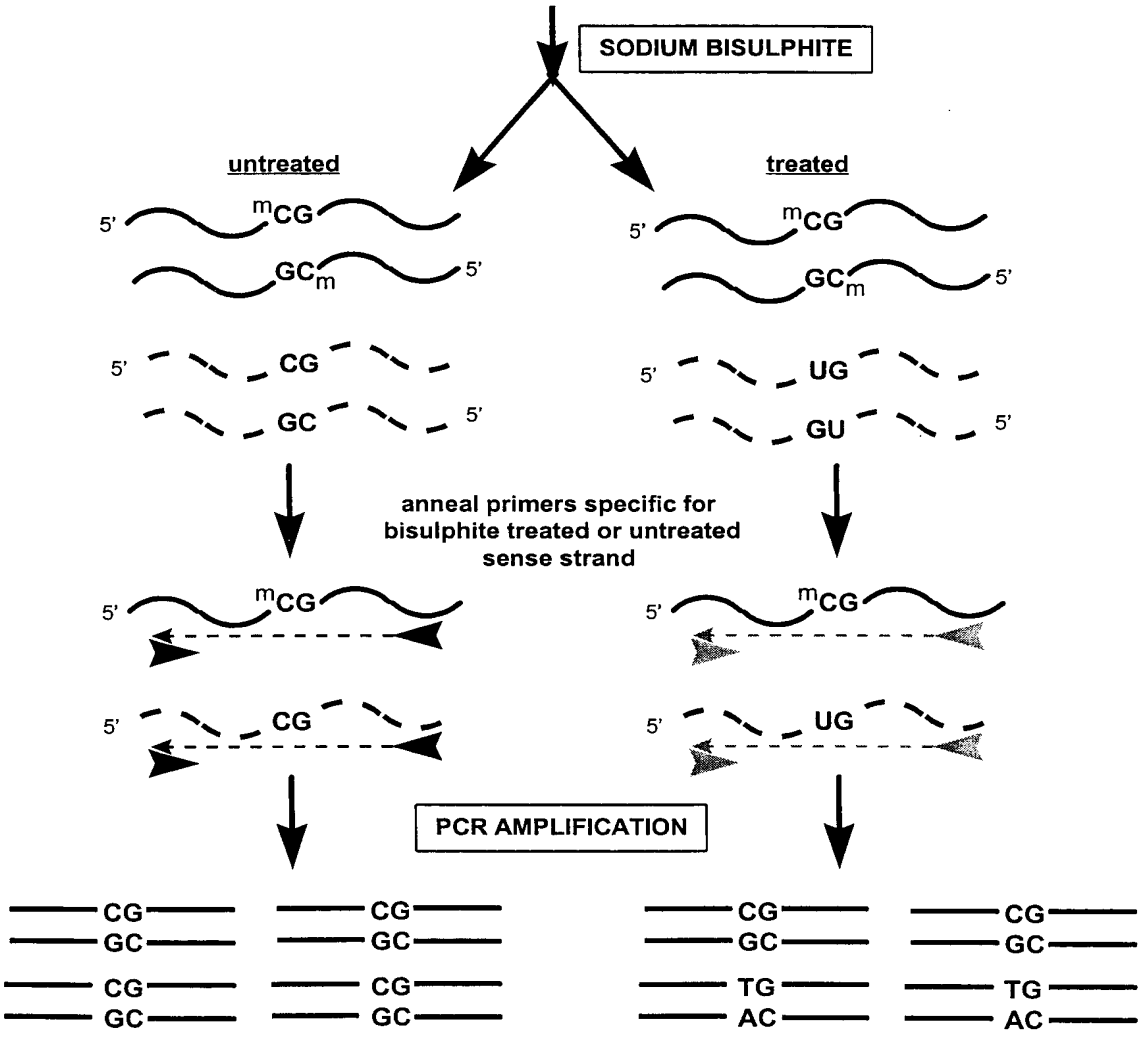
SODIUM BISULPHITE
untreated
treated
5'
mCG
GCm
5'
CG
GC
UG
GU
anneal primers specific for bisulphite treated or untreated sense strand
PCR AMPLIFICATION
TG
AC

Figure 1

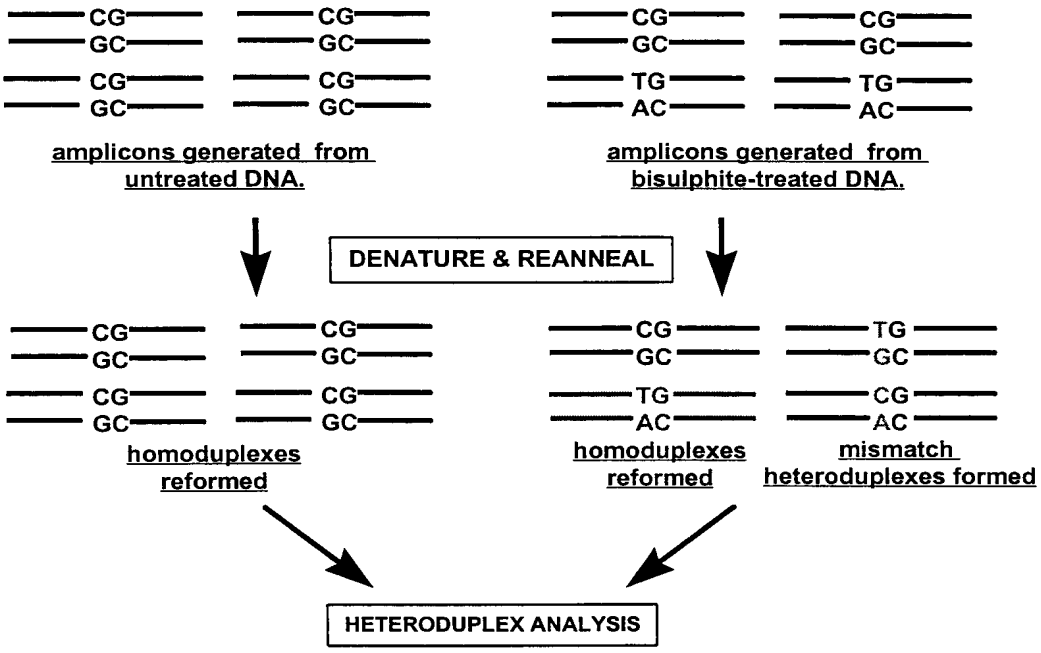
CG
GC
CG
GC
CG
GC
CG
GC
amplicons generated from untreated DNA.
CG
GC
CG
GC
TG
AC
TG
AC
amplicons generated from bisulphite-treated DNA.
DENATURE & REANNEAL
CG
GC
CG
GC
CG
GC
CG
GC
homoduplexes reformed
CG
GC
TG
GC
TG
AC
CG
AC
homoduplexes reformed
mismatch heteroduplexes formed
HETERODUPLEX ANALYSIS

Figure 2

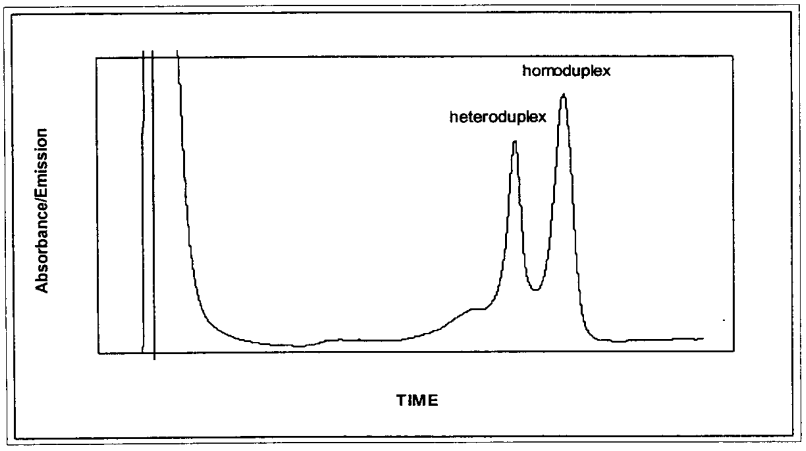

Figure 3